

Amendments to the Claims:

Please cancel Claims 1-49 and add new Claims 50-72 as follows:

Claims 1-49. (canceled).

50. (new) A method for identifying a measure of similarity between the activities of a plurality of parties, for example companies, using groups of information/text associated with, and representative of those parties on the world wide web or in other information stores, the method comprising deriving a content profile for the information group of each party, and comparing the profiles to identify a degree of similarity.

51. (new) A method as claimed in claim 50 wherein deriving the content profile of a group involves analyzing every group of text to identify key topics; allocating a measure of importance to identified key topics, and using that measure and the identified topics to generate the content profile.

52. (new) A method as claimed in claim 50 wherein the step of analyzing is based on a word frequency analysis and comprises selecting topics which have a higher than average frequency of occurrence in the group than in the native language of the group.

53. (new) A method as claimed in claim 51 wherein the step of analyzing involves discarding topics that are not related to important key words.

54. (new) A method as claimed in claim 51 comprising:
determining a list of words related to each of a plurality of key topics identified in the group; and

determining whether each key topic appears in the list of related words for any of the other key topics in the group and discarding any of the key topics where the key topics does not appear in the list of related words for any other of the key topics.

55. (new) A method as claimed in claim 51 wherein the step of comparing comprises counting the number of topics common to the profiles of each party.

56. (new) A method as claimed in claim 51 wherein comparing the profiles involves comparing the measures of importance for each key topic.

57. (new) A method as claimed in claim 51 wherein the step of comparing involves calculating an aggregated comparison across all topics common between the profiles being compared.

58. (new) A method for measuring the similarity of groups of electronic text comprising determining a content profile for each of a plurality of groups of text based electronic documents and comparing the profiles to identify a degree of similarity.

59. (new) A system for identifying a measure of similarity between the activities of a plurality of parties, for example companies, using groups of text associated with, and representative of those parties on the world wide web or in other information stores, the system being operable to derive a content profile for the information group of each party, and compare the profiles to identify a degree of similarity.

60. (new) A system as claimed in claim 59 wherein deriving the content profile of a group involves analyzing every group of text to identify key topics; allocating a measure of importance to identified key topics, and using that measure and the identified topics to generate the content profile.

61. (new) A system as claimed in claim 59 that is operable to analyze group text based on a word frequency analysis which comprises identifying key topics by selecting topics which have a higher than average frequency in the group than in the native language of the group as a whole.

62. (new) A system as claimed in claim 60 that is operable to discard topics that are not related to important key words.

63. (new) A system as claimed in claim 60 that is operable to determine a list of words related to each of a plurality of key topics identified in the group; determine whether each key topic appears in the list of related words for any of the other key topics in the group and

discard any of the key topics where the key topics does not appear in the list of related words for any other of the key topics.

64. (new) A method for profiling a group or collection of electronic text, the method comprising analyzing every group of text in the collection to identify key topics; allocating a measure of importance to identified key topics, and using that measure to generate a topic profile that includes a plurality of topic identifiers and an indication of the importance of each of the topics identified to the collection as a whole or in part.

65. (new) A method as claimed in claim 64 wherein the group of electronic document text comprises pages of a web site.

66. (new) A method as claimed in claim 64 further involving downloading each page of the site in order to do the step of analyzing.

67. (new) A method as claimed in claim 64 wherein the step of analyzing is based on a word frequency analysis which comprises identifying key topics by selecting topics which have a higher than average frequency in the group than in the native language of the group as a whole.

68. (new) A method as claimed in claim 64 wherein the step of analyzing the documents involves determining a list of words related to each of a plurality of key topics identified in the group; determining whether each key topic appears in the list of related words for any of the other key topics in the group and discarding any of the key topics where the key topics does not appear in the list of related words for any other of the key topics.

69. (new) A system for profiling a group or collection of text, the system being operable to:

analyze every document in the group of text in the collection to identify key topics; and allocate a measure of importance to identified key topics, and use that measure to generate a topic profile that includes a plurality of topic identifiers and an indication of the importance of each of the topics identified to the group as a whole.

70. (new) A system as claimed in claim 69 comprising: means for determining a list of words related to each of a plurality of key topics identified in the group; means for determining whether each key topic appears in the list of related words for any of the other key topics in the group and means for discarding any of the key topics where the key topics does not appear in the list of related words for any other of the key topics.

71. (new) A system for allowing navigation within a group of electronic documents, such as a subset of the world-wide web, the said system capable of:

automatically presenting on a screen or display a plurality of topic identifiers, together with an indication of the relative importance of the topics identified, each topic being user selectable, topics being presented in a pre-determined order, thereby to provide an indication of the importance of the topics to the group as a whole or in part; and

receiving a user selection of a given topic and providing access to information on the selected topic in response to the user selection.

72. (new) A system as claimed in claim 71, wherein said system is further capable of presenting related group identifiers for identifying one or more related groups of electronic documents, such as internet or intranet sites, together with an indication or measure of a similarity between a key topic profile of the first group and each related group.